Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Science of the Total Environment 755 (2021) 142523

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

The determinants of COVID-19 case fatality rate (CFR) in the Italian regions and provinces: An analysis of environmental, demographic, and healthcare factors

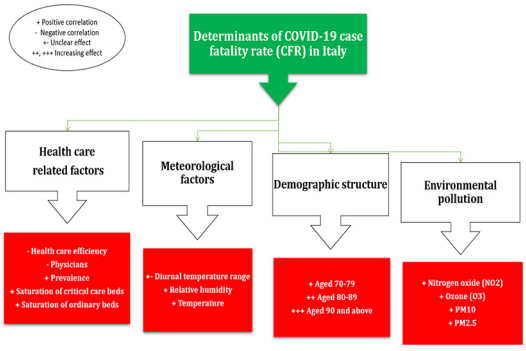
Gaetano Perone

University of Bergamo, Department of Management, Economics and Quantitative Methods, via dei Caniana 2, 24127 Bergamo, Italy

HIGHLIGHTS

- The determinants of COVID-19 CFR in the Italian regions and provinces
- Several environmental, demographic, and health system factors were studied.
- The methods used were OLS multivariate analysis and cluster analysis.
- PM_{10}, $PM_{2.5}$, NO_2, O_3, population age, humidity, and temperature were positively correlated with the CFR.
- Saturation of the health system played an important role in explaining the CFR.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 27 July 2020
Received in revised form 26 August 2020
Accepted 17 September 2020
Available online 24 September 2020

Editor: Lidia Morawska

Keywords:
COVID-19
Health system saturation
Weather
Environmental pollution
Case fatality rate
Italy

ABSTRACT

The Italian government has been one of the most responsive to COVID-2019 emergency, through the adoption of quick and increasingly stringent measures to contain the outbreak. Despite this, Italy has suffered a huge human and social cost, especially in Lombardy. The aim of this paper is dual: i) first, to investigate the reasons of the case fatality rate (CFR) differences across Italian 20 regions and 107 provinces, using a multivariate OLS regression approach; and ii) second, to build a "taxonomy" of provinces with similar mortality risk of COVID-19, by using the Ward's hierarchical agglomerative clustering method. I considered health system metrics, environmental pollution, climatic conditions, demographic variables, and three ad hoc indexes that represent the health system saturation. The results showed that overall health care efficiency, physician density, and average temperature helped to reduce the CFR. By the contrary, population aged 70 and above, car and firm density, air pollutants concentrations (NO_2, O_3, PM_{10}, and $PM_{2.5}$), relative average humidity, COVID-19 prevalence, and all three indexes of health system saturation were positively associated with the CFR. Population density, social vertical integration, and altitude were not statistically significant. In particular, the risk of dying increases with age, as 90 years old and above had a three-fold greater risk than the 80–to–89 years old and four-fold greater risk than 70–to–79 years old. Moreover, the cluster analysis showed that the highest mortality risk was concentrated in the north of the country, while the lowest risk was associated with southern provinces. Finally, since prevalence and health system saturation indexes played the most important role in explaining the CFR variability, a significant part of the latter may have been caused by the massive stress of the Italian health system.

E-mail address: gaetano.perone@unibg.it.

https://doi.org/10.1016/j.scitotenv.2020.142523
0048-9697/

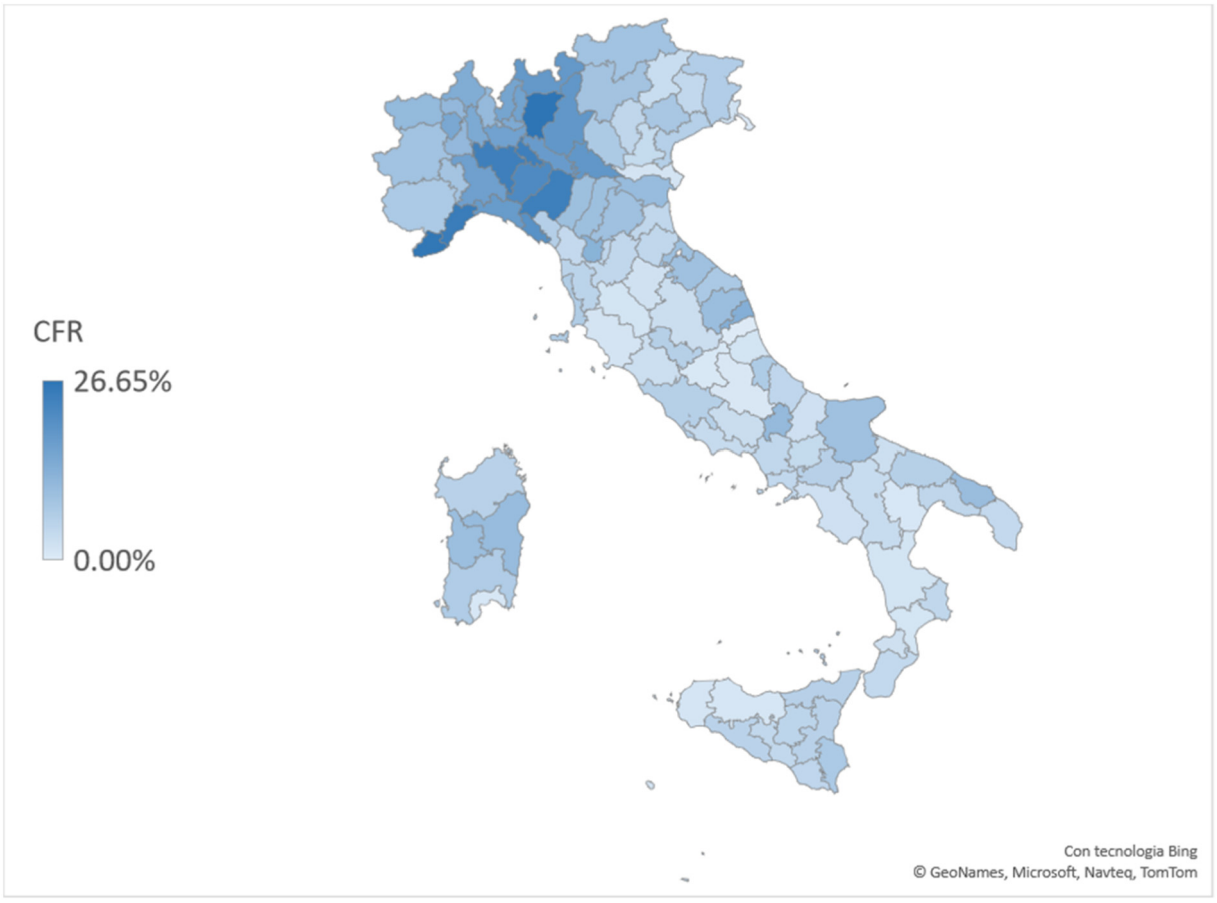


Fig. 1. The case fatality rate (CFR) at the peak of the epidemic in the Italian provinces.
Source: Italian Ministry of Health (www.salute.gov.it).

1. Introduction

The novel coronavirus disease (COVID-19) is a severe acute respiratory syndrome detected for the first time in December 2019 in Wuhan, Hubei province, China. On March 11, 2020, the World Health Organization (WHO, 2020) declared that COVID-19 could be characterized as a pandemic. As of July 26, 2020, according to Worldometer (2020), the virus has spread across 213 countries and territories, affecting over 16,4 million people and causing more than 650 thousand deaths. Italy was one of the countries hit the worst by the pandemic, with almost 250 thousand confirmed cases and over 35 thousand deaths at the time of writing. Despite the widely recognized excellence of the Italian health system (World Health Organization, 2010; GBD, 2017; Bloomberg, 2019), the country has paid a very high price, with one of the highest case fatality rate (CFR) in the world (https://ourworldindata.org/grapher/coronavirus-cfr). However, at the peak of the epidemic, mortality was affected by a large spatial heterogeneity; in fact, the northern regions were characterized on average by a significantly higher CFR than southern regions (Fig. 1).

The aim of this work is to contribute to the existing literature by investigating the main reasons and determinants of COVID-19 CFR at the peak of the virus outbreak in 20 Italian regions and 107 Italian provinces. To reach this goal, I considered several variables: health system metrics, air pollutants concentrations, climatic conditions, demographic factors, and three ad hoc indexes that represent the health system saturation.

Specifically, I chose the peak of mild and severe cases of COVID-19, which was approximately reached on April 3–4, 2020 (Fig. 2), because it can be considered as the moment of maximum health care system saturation. This is an important aspect, because at the early outbreak stages the number of infected usually growths exponentially and health system cannot systematize its response. So, the health system saturation may assume greater importance.[1]

2. Relevant literature

In the last months, a plenty and increasing body of literature focused its attention on the environmental, meteorological, demographic, and social factors that may affect COVID-19 mortality (Bayer and Kuhn, 2020; Brandt et al., 2020; Comunian et al., 2020; Du et al., 2020; Ma et al., 2020; Pansini and Fornacca, 2020; Sannigrahi et al., 2020; X. Wu et al., 2020; Verity et al., 2020; Zhu et al., 2020). What follows is a summary of the main findings on the relationship between COVID-19 CFR and climatic conditions, demographic variables, and air pollutants concentrations, respectively.

Ma et al. (2020) used a generalized additive model (GAM) to study the relationship between the meteorological factors and the daily deaths of COVID-19 in Wuhan from January 20, 2020 to February 29, 2020. They found a positive association between the daily deaths of COVID-19 and the diurnal temperature range (DTR), and a negative relationship between the former and the relative humidity and temperature. Similarly, Y. Wu et al. (2020) used a log-linear GAM to analyze the effect of temperature and humidity on daily new deaths of COVID-19 in 166 countries, as of March 27, 2020. Temperature and relative humidity

[1] In other words, I assumed that in the early stage of COVID-19, the capacity of health system to respond to the rapid increase in the demand for hospital beds was rather limited.

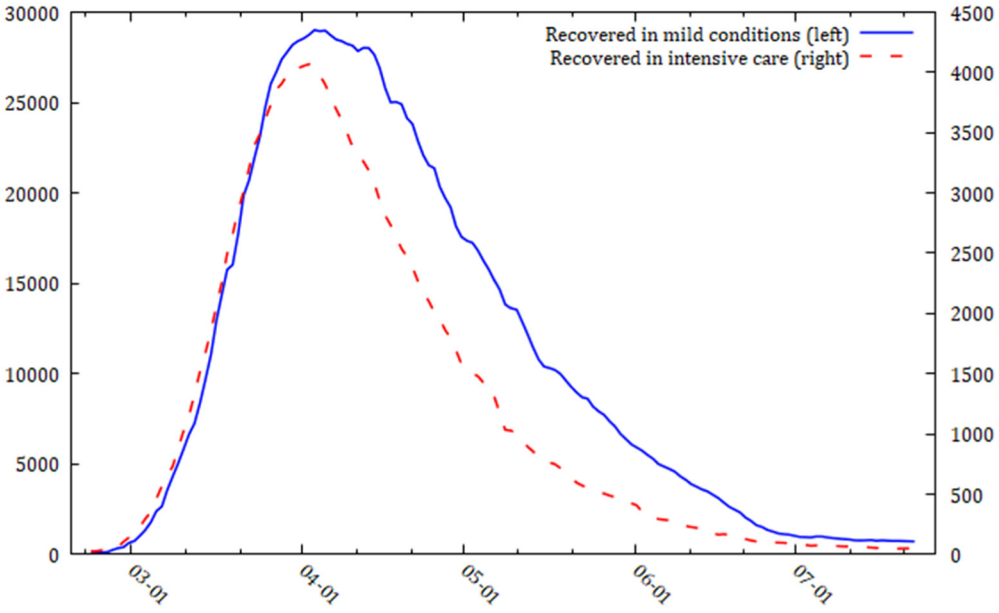


Fig. 2. The number of people recovered from COVID-19 in the period February 23, 2020–July 23, 2020.
Source: Italian Ministry of Health (www.salute.gov.it).

were found to be both significantly and negatively associated with daily new deaths. Rahman et al. (2020) performed partial correlation analysis and linear mixed effect modeling to analyze the effect of temperature on COVID-19 mortality risk in 149 countries. They showed that higher temperatures were negatively associated with mortality in high-income countries, while extreme temperature may increase mortality risk in low-and middle-income countries.

Sannigrahi et al. (2020) used spatial regression models to analyze the spatial association between the key demographic variables and COVID-19 deaths across 31 European countries. They found that the incidence of the population aged 80 and above on overall casualties caused by COVID-19 was highly significant. Verity et al. (2020) used a model-based approach to estimate the case fatality ratio associated with age groups in mainland China. They found that the population aged 80 and above had the highest case fatality ratio (13.4%). Similarly, Du et al. (2020), by implementing a univariate and multivariate logistic regression analysis, showed that advanced age is a significant risk factor for COVID-19 mortality. Moreover, Ioannidis et al. (2020), by using official data from 14 countries and 13 US states as of June 17, 2020, estimated that age risk gradients were highly significant. In particular, people aged less than 65 had 16–100-fold lower risk of COVID-19 deaths than older people.

In early March 2020, Bayer and Kuhn (2020) advanced the hypothesis that Italian higher vertical social integration may arise the COVID-19 fatality rate. They used a sample of 24 countries with at least 200 COVID-19 confirmed cases and found a positive correlation between the share of the population aged 30–49 living with parents and COVID-19 death rate. However, this relation was strongly criticized by Belloc et al. (2020), which showed how the sign of the correlation turned negative when considering the variation within Italian regions.

Regarding the effect of air pollutants concentrations on COVID-19 related deaths, Conticini et al. (2020), by investigating the relevant literature, concluded that prolonged exposure to air pollutants may lead to chronic respiratory conditions, even in healthy and young people. X. Wu et al. (2020) analyzed COVID-19 death counts for 3087 counties in the USA, covering 98% of the population, by using a negative binomial mixed model. They found that a positive and significant association between $PM_{2.5}$ and COVID-19 mortality rates. In particular, a 1 unit increase in $PM_{2.5}$ is related to an 8% increase in the COVID-19 fatality rate. Pansini and Fornacca (2020), using Kendall's tau and Pearson correlation coefficient, found a positive and significant association between COVID-19 mortality and several air pollutants (CO, NO_2, PM_{10}, $PM_{2.5}$) in China and the USA.

Ogen (2020) used spatial analysis to examine the relationship between long-term exposure to NO_2 and COVID-19 mortality in 66 administrative regions in France, Germany, Italy, and Spain. He showed that 78% of the total COVID-19 deaths were located in north Italy and central Spain, i.e. the regions with the highest level of NO_2. Bianconi et al. (2020), using multiple linear regression models, showed that mean annual exposure to PM_{10} and $PM_{2.5}$ was significantly and positively associated with COVID-19 death rate in Italian 20 regions. Yao et al. (2020), by using spatial analysis and multivariate linear regression for China as of April 12, 2020, found that every 10 $\mu g/m^3$ increase in PM_{10} and $PM_{2.5}$ concentrations is associated with a 0.24% and 0.26% increase in the COVID-19 mortality rate, respectively. Brandt et al. (2020) also stressed that air pollution levels are strongly associated with densely populated urban areas in the USA. Hamidi et al. (2020) investigated both direct and indirect impacts of population density on death rates in 913 USA metropolitan countries by using structural equation modeling. They found that larger metropolitan areas and counties with higher population density were significantly associated with higher mortality rates.

Finally, several studies have shown that the presence of at least one comorbidity, such as hypertension, diabetes, cardiovascular disease, and chronic lung disease, negatively affects outcomes of patients hospitalized with COVID-19 (Guan et al., 2020; Istat-ISS, 2020; Jordan et al., 2020; Wang et al., 2020; Wu and McGoogan, 2020).

3. Material

At regional level, I used the following 17 explanatory variables (Table 1): an overall index (IPS) of the Italian health system performance in 2017–2018, the public health expenditure per capita in the period 2015–2017, the total specialist doctors and general practitioners per 1000 inhabitants in the period 2016–2018, the total ordinary hospital beds per 1000 inhabitants in the period 2016–2018, an index of car and firm density in 2015–2017, the

Table 1
Definitions of all variables used for OLS regional analysis.

Variables	Definitions	Sources
Dependent variables		
CFR	The average case fatality rate for COVID-19 in each region, obtained by dividing the average confirmed deaths by the average confirmed cases on April 3 and 4, 2020.	Italian Ministry of Health[a]
Independent variables		
IPS	A synthetic index of the Italian health system performance in the period 2017–2018, which includes eight different parameters.[b]	Demoskopika Research Institute (2018, 2019)
Health expenditure	The average public health expenditure per capita for each region, in the period 2015–2017.	I.Stat (database)[c]
Physicians	The average total specialist doctors and general practitioners (per 1000 inhabitants) for each region, in the period 2016–2018.	I.Stat (database)
Hospital beds	The average ordinary hospital beds (per 1000 inhabitants) for each region, in 2016–2018.	Italian Ministry of Health
Cars & Firms	A synthetic index of car and firm (> 250 employees) density for each region, in 2015–2017.[d]	I.Stat (database)
kWh per capita	The average electric power consumption in kilowatt-hours (kWh) per capita for each region, in the period 2016–2018.	Terna (2019)
Ages 70+	The proportion of population aged 70 and over for each region, in 2019.	I.Stat (database)
Ages 80+	The proportion of population aged 80 and over for each region, in 2019.	I.Stat (database)
Ages 90+	The proportion of population aged 90 and over for each region, in 2019.	I.Stat (database)
Vertical integration	The share of unmarried young adults aged 18–34 living with at least one parent for each region, in 2019.	I.Stat (database)
Humidity	The average relative humidity levels registered during March 2020, for each region.[e]	www.il meteo.it[f]
DTR	The historical diurnal temperature range in March, for each region.	Mipaaf (2019a)
Temperature	The historical average temperature in March, for each region.	Mipaaf (2019a)
Prevalence	The average ratio between the people who have been tested positive for COVID-19 and the overall population of each region on April 3 and 4, 2020.	I.Stat (database), Italian Ministry of Health
Preval./Beds	The ratio between the average COVID-19 prevalence on April 3 and 4, 2020, and the average number of ordinary hospital beds in 2016–2018, for each region.	Italian Ministry of Health
CCB saturation	The ratio between the average people who have been recovered from COVID-19 in intensive care on 3 and 4 April 2020, and the average number of critical care beds (CCB) in the period 2016–2018, for each region.	Italian Ministry of Health
OB saturation	The ratio between the average people who have been recovered from COVID-19 with mild symptoms on 3 and 4 April 2020, and the average number of ordinary hospital beds in the period 2016–2018 for each region.	Italian Ministry of Health

[a] Data are available at URL: www.salute.gov.it.
[b] The parameters used are the following: patient satisfaction, active patient mobility, passive patient mobility, legal fees for disputes, operating result, life expectancy, equality in health treatment, and economic hardship. In particular, each parameter is standardized, with mean = 100 and standard deviation = 10, and the final synthetic index is obtained by calculating the simple average of them.
[c] Data are available at URL: http://dati.istat.it/.
[d] The number of cars refers to those recorded in the Pubblico registro automobilistico (Public vehicle register). The number of largest firms (> 250 employees) refers to those that operate in the following sectors: (1) mining and minerals from quarries and mines; (2) manufacturing activities; (3) supply of electricity, gas, vapors, and air conditioning; and (4) supply sewerage, waste management and remediation activities. The index is compiled according to the following analytical method: i) first, I standardized the data according to surface area (cars and firms for 100 sq. km.); ii) than, the respective outputs are switched to fixed-base indexes (with mean = 100); iii) finally, I computed the simple arithmetic mean of the latter.
[e] The average values have been calculated by dividing the data coming from 62 different official weather stations, managed by the Italian Air force and located in the main Italian provinces.
[f] This is one of the most trusted Italian weather forecast website. https://www.ilmeteo.it/business/assets//images/aboutUs/pdf/Google%2009_04_2020.pdf.

electric power consumption (kWh per capita) in the period 2016–2018,[2] the proportion of population aged 70 and over in 2019, the proportion of population aged 80 and over in 2019, the proportion of population aged 90 and over in 2019, the vertical social integration proxied by the share of adults aged 18–34 living with their parents in 2019, the average relative humidity levels registered during March 2020, the average historical diurnal temperature range (DTR) in March, the average historical temperature in March, the average prevalence of COVID-19 on April 3 and 4, 2020, the ratio between the COVID-19 prevalence and the ordinary hospital beds, an two ad hoc indexes that indicate the saturation of ordinary hospital beds (OB) and critical care beds (CCB) at the peak of the outbreak, respectively. The dependent variable is represented by the average case fatality rate of COVID-19 on April 3 and 4, 2020.[3]

At the province level, I used the following 18 explanatory variables (Table 2): the average general practitioners per 1000 inhabitants in the period 2015–2017, the historical average temperature in March, the historical average diurnal temperature range (DTR) in March, an ordinal index of the population structure (rural-intermediate-urban) in 2013, the population density (people per sq. km) in 2019, the proportion of population aged 70 and over in 2019, the proportion of population aged 70–79 in 2019, the proportion of population aged 80–89 in 2019, the proportion of population aged 90 and over in 2019, the average concentrations (in $\mu g/m^3$) of particulate matter less than 10 μm in diameter (PM_{10}) in 2017–2018, the number of days in which PM_{10} concentrations exceeded the legal limit of 50 $\mu g/m^3$,[4] the average concentrations (in $\mu g/m^3$) of particulate matter less than 2.5 μm in diameter ($PM_{2.5}$) in the period 2017–2018, the average concentrations of nitrogen dioxide (NO_2) expressed in $\mu g/m^3$ in 2017–2018, the number of days in which ozone (O_3) concentrations exceeded the limit of 120 $\mu g/m^3$,[5] the average altitude of the capital city of each province, the prevalence of COVID-19 on March 31, 2020, and the ratio between the COVID-19 prevalence on March 31 and the average number of ordinary hospital beds in the period 2016–2018. The dependent

[2] Car and firm density and electricity consumption can be considered as two proxy indicators for air pollution (Gan et al., 2012; Tian et al., 2007).
[3] According to Baud et al. (2020) and Scheiner et al. (2020), there is a delay between infection and death of about 14 days. However, on one hand, the prevalence on March 19–21 and April 3–4 had an almost perfectly positive correlation of 0.96 (elaboration on data from http://www.salute.gov.it),and on the other, the CFR increased exponentially until the peak, and then it has continued to grow much more slowly (https://ourworldindata.org/grapher/coronavirus-cfr).
[4] The legal limit is laid down in Directive 2008/50/EC (https://eur-lex.europa.eu/legal-content/en/ALL/?uri=CELEX%3A32008L0050).
[5] The average concentrations of air pollutants for each province have been calculated by using the annual data coming from a maximum of 266 urban and suburban monitoring stations, spread all over the country (Istat, 2019).

Table 2
Definitions of all variables used for OLS province analysis.

Variables	Definitions	Sources
Dependent variables		
Death rate	The average case fatality rate for COVID-19 in each province, obtained by dividing the confirmed deaths by the number of confirmed cases, on 31 March 2020.	Istat-ISS (2020)
Independent variables		
General practitioners	The average general practitioners for each province, in 2019.	Il Sole 24 Ore (2019)
Temperature	The historical average temperature in March, for each province.	Mipaaf (2019b)
DTR	The historical average diurnal temperature range in March, for each province.	Mipaaf (2019b)
Urbanization	An ordinal index that ranks population of each province by urban-rural structure: predominantly rural (1), intermediate (2), and predominantly urban (3).	European Commission (2013)
Density	The number of human inhabitants per square kilometer (sq. km.) of land area for each province, in 2019.	I.Stat (database)
Ages 70+	The proportion of population aged 70 and over for each province, in 2019.	I.Stat (database)
Ages 70–79	The proportion of population aged 70–79 for each province, in 2019.	I.Stat (database)
Ages 80–89	The proportion of population aged 80–89 for each province, in 2019.	I.Stat (database)
Ages 90+	The proportion of population aged 90 and over for each province, in 2019.	I.Stat (database)
PM_{10}	The average concentrations of particulate matter less than 10 μm in diameter, expressed in $\mu g/m^3$ for each province, in the period 2017–2018.	Istat (2019)
PM_{10} (>50)	The average number of days in which PM_{10} exceeded the limit of 50 $\mu g/m^3$ for each province, in the period 2017–2018.	Istat (2019)
$PM_{2.5}$	The average concentrations of particulate matter less than 2.5 μm in diameter, expressed in $\mu g/m^3$ for each province, in the period 2017–2018.	Istat (2019)
NO_2	The average concentrations of nitrogen dioxide, expressed in $\mu g/m^3$ for each province, in the period 2017–2018.	Istat (2019)
O_3	The average number of days in which ozone exceeded the limit of 120 $\mu g/m^3$ for each province, in the period 2017–2018.	Istat (2019)
Altitude	The average altitude of the capital city of each province.	Istat (2019)
Prevalence	The ratio between the people who have been tested positive for COVID-19 on March 31, 2020, and the total population of each province in 2019	I.Stat (database), Italian Ministry of Health
OB saturation	The ratio between the COVID-19 prevalence on March 31, 2020, and the average number of ordinary hospital beds in 2016–2018, for each province.	Italian Ministry of Health

variable is represented by the case fatality rate of COVID-19 on March 31, 2020.[6] Further details are provided in Tables 1 and 2.[7]

4. Methods

First, I used a multivariate cross-sectional OLS (ordinary least squares) approach to identify the main determinants at regional at province level, and then I applied the Ward's hierarchical agglomerative clustering method (Ward Jr, 1963) to build a "taxonomy" of provinces with similar mortality risk of COVID-19. The rationale behind the choice of cross-sectional regression methodology instead of a panel approach is as follows: i) comprehensive daily data on March 2020 are not currently available for air pollutants concentrations and climatic variables; and ii) demographic and health system variables change very little or not at all in the short term.

The OLS equation estimated at the regional and provincial level was as follows:

$$y_i = \beta_0 + \beta_1 X_1 + \ldots \beta_n X_n + \varepsilon_i \quad (1)$$

where *i* represented the regions/provinces, β_0 was the intercept, X_1 to X_n were the independent variables for each region/province, and ε_i was the random error.

Then, I conducted a cluster analysis on selected significant variables obtained from multivariate OLS on province data. The procedure encompassed the following 4 sequential steps. First, since variables are not on the same scale, I calculated the standardized data matrix by using the following formula:

$$z = \frac{x-\mu}{\sigma} \quad (2)$$

where x was the value of the variable in the original dataset, μ was the arithmetic mean of the original variable, and σ was the standard deviation of the latter. Second, I computed the Euclidean distance. Given two points X and Y in d dimensional space, the Euclidean distance between X and Y was equal to:

$$\|X-Y\| = \sqrt{\sum_{i=1}^{d}(x_i-y_i)^2} \quad (3)$$

Then, I applied Ward's hierarchical clustering method, which allowed to obtain clusters of provinces with features as similar as possible by minimizing the total within-cluster variance. Specifically, at each step, the pair of clusters that are characterized by minimum between-cluster distance were merged. Therefore, Ward's method merging cost formula between two clusters, p and q, was given by:

$$\Delta(p,q) = \sum_{i\in p\cup q}\|x_i - m_{p\cup q}\|^2 - \sum_{i\in p}\|x_i - m_p\|^2 - \sum_{i\in q}\|x_i - m_q\|^2 \quad (4)$$

From this, it was obtained:

$$\Delta(p,q) = \frac{n_p n_q}{n_p + n_q}\|m_p - m_q\|^2 \quad (5)$$

where m_j was the center of cluster *j*, and n_j was the overall number of points included in cluster *j*.

5. Results and discussion

5.1. OLS at the regional level

In Tables 3a, 3b), I presented the OLS estimations at the regional level. All the 12 estimated OLS models were statistically significant; in fact, the Fisher-Snedecor distribution assumed values far higher than the tabulated critical values at the 1% level of significance. In particular, the r-square showed that models were able to explain from 0.44% to 0.88% of CFR variability. Furthermore, since Breusch and Pagan (1979) and Shapiro and Wilk (1965) tests allowed to accept the null hypothesis of homoscedasticity and normality of residuals, models seemed well specified. However, due to the small sample, I preferred to adopt a

[6] Istat released only the overall death for March 31, April 30, and May 31, 2020.
[7] Some descriptive statistics are also reported in the Supplementary Material (Tables S1 and S2).

Table 3a
(Models 1–6). OLS regression at the regional level between CFR and environmental, demographic, and healthcare factors.

Variables	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Constant	16.4788	21.4019	27.8448*	−12.1666	−11.7866	−14.8309
	[13.9432]	[12.8697]	[12.7974]	[28.5844]	[12.8946]	[13.3257]
IPS	−0.4788*	−0.5451**	−0.5731**	−0.5758**	−0.4511**	−0.4808***
	[0.2303]	[0.2221]	[0.2153]	[0.2034]	[0.1402]	[0.1468]
Health exp.	0.0129**	0.0157**	0.0144**	0.0237**	0.0146*	0.0168**
	[0.0059]	[0.0056]	[0.0058]	[0.0096]	[0.0074]	[0.0067]
Physicians	−3.2988**	−3.1879**	−3.1683**	−2.3358	−3.0878**	−2.5227*
	[1.2894]	[1.1264]	[1.136]	[1.8958]	[1.3113]	[1.1406]
H. Beds	3.8867	2.7984	3.2841	−1.0369	−1.4253	−2.1594
	[2.8595]	[2.8415]	[3.1279]	[5.1024]	[1.7597]	[2.191]
Car & Firm	6.2449***	7.0056***	6.9565***	8.5611***	4.5281**	5.4669***
	[1.7936]	[1.7842]	[1.93]	[1.9709]	[1.4903]	[1.4142]
Kilowatt	−0.0007	−0.0005	−0.0005	−0.0015*	−0.0009	−0.0012*
	[0.0006]	[0.0006]	[0.0007]	[0.0007]	[0.0007]	[0.0006]
Aged 70+	0.7821**			1.0916**	0.7973***	0.8982***
	[0.3183]			[0.4199]	[0.2343]	[0.234]
Aged 80+		1.5**				
		[0.5131]				
Aged 90+			6.3807**			
			[2.5141]			
Humidity				0.449**	0.1413	0.2615**
				[0.1886]	[0.1535]	[0.0808]
DTR				0.3657	2.679***	2.3164***
				[1.5456]	[0.6612]	[0.6588]
Temperature				−1.1144*	0.5302	
				[0.5752]	[0.4919]	
Prevalence					26.5164***	22.0411***
					[4.3121]	[5.3084]
Breusch-P. (p)	0.2454	0.5101	0.3985	0.7505	0.4038	0.6391
Shapiro-W. (p	0.9892	0.9969	0.9713	0.7997	0.0115	0.1289
Influential (h)	0.13–0.78	0.15–0.74	0.17–0.69	0.23–0.82	0.25–0.87	0.2–0.87
VIF	1.56–4.44	1.46–4.45	1.49–4.88	2.49–5.43	2.65–8.28	2.04–5.79
F-statistic	5.85***	5.28***	5.5***	5.18***	22.13***	22.55***
Observations	20	20	20	20	20	20
Adjusted R^2	0.4425	0.4753	0.4531	0.4899	0.8102	0.8127

Notes: h, leverage; p, p-value. Standard errors (in brackets) are based on HC2 method developed by MacKinnon and White (1985). Significance level: p-value <0.01***; p-value <0.05**: p-value <0.1*.

conservative approach, by applying the HC2 correction proposed by MacKinnon and White (1985).[8] It is important to stress that OLS cross-sectional analysis is very sensitive to the presence of outliers, which can cause misspecification issues (Mur and Lauridsen, 2007), especially in such small samples (Wooldridge, 2015, p. 334). Therefore, I investigated the presence of highly influential points. The analysis revealed that none of the leverage points (h) are beyond the cutoff value ($h > 2n/k$) proposed by Belsley et al. (1980). Finally, I also identified the presence of multicollinearity. The variance inflation factors (VIF) ranged from 1.46 to 8.28[9] and were less than 10, i.e. the rule of thumb suggested by the relevant literature (Belsley, 1982; Hair Jr. et al., 1995). Therefore, I concluded that the independent variables just suffered from weak linear dependency and there were no serious multicollinearity issues.

The results showed that the IPS index and physicians were significantly and negatively associated with COVID-19 CFR.[10] By the contrary, health expenditure, car and firm density, people aged 70 and above, humidity, DTR, prevalence, the ratio of prevalence/ordinary beds, and saturation indexes of ordinary and critical care beds were significantly and positively correlated with the CFR. Hospital beds, kWh per capita, and social vertical integration showed no statistical significance. Even if the sign of health expenditure is unexpected, the IPS index seemed to confirm the importance of health system effectiveness in all its dimensions more than only health expenditure. The regression coefficients also gave interesting information. The change in 1 unit of physician per 1000 inhabitants and IPS index was correlated on average with a change of −2.87% and − 0.53% in the CFR, respectively.[11] Consistent with Du et al. (2020), Ioannidis et al. (2020), Sannigrahi et al. (2020), the risk of dying increased progressively with the increase in the elderly population. In fact, the regression coefficients of the population aged 90 and above were 4 times larger than that for the population aged 80 and above, and 8 times larger than that for the population aged 70 and above.

Moreover, each 1 unit increase in humidity and DTR was positively associated on average with a change of 0.28%[12] and 2.22% in the CFR. The results for humidity are consistent with a recent study by Bianconi et al. (2020) on the Italian case, while DTR outcomes are in line with Ma et al. (2020). Prevalence and health system saturation indexes were highly significant in all the models. In particular, saturation of ordinary beds had the largest coefficient, followed by COVID-19 prevalence, the ratio prevalence/beds, and the saturation of intensive care beds. They played the most important role in explaining CFR variability; in fact, the average r-square of models (5–12) with prevalence and saturation indexes was 0.81, i.e. almost twice the models (1–4) without them ($r = 0.46$). In particular, the change in 0.1 unit of CCB and OB

[8] As suggested by Long and Ervin (2000, p. 220), in presence of homoskedasticity, HC2 option has good small size properties.

[9] I did not consider VIF for average temperature, which was constantly greater than 10. It is due to this that average temperature has been excluded in most part of the models.

[10] In addition, in four OLS models, the average electric power consumption is negatively associated with the CFR. However, its coefficients are statistically significant only at the 10% level.

[11] From now on, I only considered the average of the significant coefficients.

[12] Model 10 also seemed to suggest that the relationship between humidity and CFR is non-monotonic. In other words, the CFR decreased for low-medium levels of humidity and increased for high values of humidity. However, since the lack of comprehensive data on humidity, further investigations are necessary.

Table 3b
(Models 7–12). OLS regression at the regional level between CFR and environmental, demographic, and healthcare factors.

Variables	Model 7	Model 8	Model 9	Model 10	Model 11	Model 12
Constant	−9.9732	−3.5934	−12.2744	168.707**	−4.8115	8.2251
	[12.8992]	[12.9892]	[13.7815]	[65.3856]	[20.9305]	[12.6372]
IPS	−0.5065***	−0.5475***	−0.4995***	−0.671***	−0.6399***	−0.4174**
	[0.1497]	[0.1504]	[0.1506]	[0.1663]	[0.1731]	[0.1348]
Health exp.	0.0151**	0.0155**	0.0163**	0.012**	0.0156	0.0062
	[0.0056]	[0.0056]	[0.0069]	[0.0051]	[0.0099]	[0.0074]
Physicians	−2.4044*	−1.8737	−2.4275*	−1.8232	−1.838	−0.8314
	[1.2974]	[1.2409]	[1.1621]	[1.0917]	[1.3513]	[0.9748]
H. Beds			−0.7423	−1.2938	1.3549	−1.0043
			[2.3357]	[1.4252]	[3.2707]	[2.4216]
Car & Firms	5.2252***	5.7172***	5.395***	7.2545***	6.4978***	0.6217
	[1.391]	[1.4174]	[1.4313]	[1.2172]	[1.4789]	[1.6553]
Kilowatt	−0.0012	−0.0009	−0.0012*	−0.001*	−0.0007	0.0000
	[0.0007]	[0.0007]	[0.0006]	[0.0004]	[0.0006]	[0.0004]
Aged 70+	0.7931**		0.7943**		0.8673**	0.2744
	[0.263]		[0.2501]		[0.3488]	[0.3046]
Aged 80+		1.3558**		2.1616***		
		[0.4649]		[0.6187]		
Vertical Integ.	−0.008	0.0254				
	[0.1366]	[0.1343]				
Humidity	0.2455**	0.2108**	0.2536**	−4.6525**	0.2824**	−0.0854
	[0.0956]	[0.0927]	[0.084]	[1.8055]	[0.1213]	[0.0711]
Humidity^2				0.036**		
				[0.0134]		
DTR	1.9813**	1.6247**	2.0091**	2.5458***	1.18	2.3931**
	[0.7199]	[0.6522]	[0.6946]	[0.6437]	[1.0982]	[0.8386]
Preval./beds	7.2081***	7.121***	7.3835***	6.9567***		
	[2.0728]	[2.0874]	[1.8403]	[1.1719]		
ICB saturation					4.6762**	
					[1.5899]	
OB saturation						45.4811***
						[7.9362]
Breusch-P. (p)	0.5386	0.6034	0.6502	0.5497	0.8678	0.645
Shapiro-W (p)	0.087	0.8456	0.0951	0.9033	0.1169	0.6978
Influential (h)	0.2–0.85	0.21–0.85	0.21–0.87	0.23–0.92	0.2–0.81	0.2–0.81
VIF	2.32–5.89	2.08–5.91	1.95–5.43	–	2.31–5.48	2.03–7.25
F-statistic	18.96***	25.07***	21.32***	20.24***	9.57***	54.14***
Observations	20	20	20	20	20	20
Adjusted R^2	0.7965	0.799	0.7982	0.8782	0.7024	0.8576

Notes: h, leverage; p, p-value. Standard errors (in brackets) are based on HC2 method developed by MacKinnon and White (1985). Significance level: p-value <0.01***; p-value <0.05**: p-value <0.1*.

saturation was correlated with a change of 0.47% and 4.54% in the CFR. Therefore, a significant part of the CFR may be caused by the massive stress of the Italian health system.

Finally, the correlation matrix in Fig. S1 indicated that the average temperature and humidity were significantly and negatively correlated to the COVID-19 prevalence, with a Pearson's r of −0.8 and −0.46, respectively. By the contrary, the kWh per capita and car and firm density showed a good positive correlation with the latter, with a Pearson's r of 0.66 and 0.52, respectively.

5.2. OLS at the provincial level

In Tables 4a, 4b, I presented the OLS estimations at the provincial level. All 11 estimated OLS models were statistically significant; in fact, the Fisher-Snedecor distribution assumed values far higher than the tabulated critical values at 1% level of significance. In particular, r-square showed that models were able to explain from 0.27% to 0.45% of CFR variability.

Since the Breusch-Pagan (1979) test revealed heteroscedasticity issues, I applied the HC2 correction. As stated by Ghasemi and Zahediasl (2012), the violation of normality assumption should not be a major problem in a sample with enough observations ($n > 40$), such as in this case. Furthermore, the variance inflation factors (VIF) ranged from 1.05 to 2.56 and are much less than 10; thus, there were no multicollinearity issues.

The OLS models showed that general practitioners, average temperature, and DTR were significantly and negatively associated with the CFR.[13] By the contrary, all the considered air pollutants (PM_{10}, $PM_{2.5}$, NO_2, and O_3), people aged 70 and above, prevalence, and saturation index for ordinary beds were significantly and positively correlated with the CFR. Urbanization, population density, and altitude showed no statistical significance. The change in 1 degree Celsius of temperature and DTR was correlated with an average change of −0.53% and −2.44% in the CFR, respectively. If the results for the temperature are consistent with other studies (Bianconi et al., 2020; Rahman et al., 2020; Y. Wu et al., 2020), the sign of DTR is unexpected and in contrast with previous findings. Therefore, the relative impact is ambiguous and unclear.

Among air pollutants, $PM_{2.5}$ and PM_{10} had the largest impact on COVID-19 CFR, followed by NO_2 and O_3. The change in 10 μg/m^3 of $PM_{2.5}$ and PM_{10} was associated with a change of 3.13% and 2.54% in the CFR, respectively. The change in 10 μg/m^3 of NO_2 was "only" correlated with a change of 1.27% in the CFR. While, each day in which O_3 exceeded the limit of 120 μg/m^3, the CFR have increased of 0.11%. The inclusion of O_3 is very significant, since it allowed r-square to increase to 0.42. The positive and significant relationship between air pollutants and CFR is consistent with the relevant literature (Bianconi et al., 2020; Ogen, 2020; Pansini and Fornacca, 2020; X. Wu et al., 2020; Yao et al., 2020).

[13] It is important to stress that signs for temperature and DTR didn't change even considering the historical values of winter. This was also true for the regional analysis.

Table 4a
(Models 1–6). OLS regression at the provincial level between CFR and environmental, demographic, and healthcare factors.

Variables	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Constant	21.3371*	22.645**	21.4665**	28.7773***	27.8601**	28.3833**
	[10.8017]	[11.0182]	[10.4233]	[10.0096]	[11.075]	[11.7957]
G. P.	−10.0768**	−9.4475**	−10.7185**	−10.8571**	−9.9708**	−12.8379***
	[4.2187]	[4.1747]	[4.298]	[4.2279]	[4.0876]	[4.2603]
Temperature	−0.5506**	−0.5966**	−0.5292**	−0.5296**	−0.4515*	−0.4038
	[0.242]	[0.236]	[0.2431]	[0.2449]	[0.2455]	[0.255]
DTR	−2.5289**	−2.6147**	−2.3998**	−2.6982**	−2.516**	−2.5241**
	[1.0323]	[1.0539]	[1.0138]	[1.0269]	[1.0833]	[1.049]
PM_{10} (μg/m^3)	0.2569***	0.2529***	0.256***	0.252***		
	[0.0939]	[0.0952]	[0.0938]	[0.0931]		
Urbanization	0.8979	0.7571	1.0869	1.0815		
	[0.9731]	[0.9805]	[0.97]	[0.9602]		
Pop. Density	−0.0004	−0.0005	−0.0004	−0.0003		
	[0.0012]	[0.0012]	[0.0012]	[0.0012]		
Aged 70+	0.7596**				0.6823**	0.6955**
	[0.3106]				[0.3157]	[0.3149]
Aged 70–79		1.2866**				
		[0.6163]				
Aged 80–89			1.9984***			
			[0.7511]			
Aged 90+				5.5753**		
				[2.2503]		
PM_{10} (> 50)					0.0739***	
					[0.0257]	
NO_2 (μg/m^3)						0.1268**
						[0.055]
Breusch-P p.	0.0334	0.05	0.043	0.0385	0.0049	0.0083
F-statistic	6.37***	5.9***	6.33***	6.95***	9.81***	9.0808***
VIF	1.11–1.56	1.09–1.56	1.12–1.55	1.18–1.57	1.06–1.43	1.07–1.52
Observations	107	107	107	107	107	107
Adjusted R^2	0.2944	0.2775	0.306	0.2903	0.301	0.266

Notes: p, p-value. Standard errors (in brackets) are based on HC2 method developed by MacKinnon and White (1985). Significance level: p-value <0.01***; p-value <0.05**: p-value <0.1*.

As found at the regional level, the CFR increased progressively with population age. Specifically, 90 years old and above had a three-fold greater risk than the 80–to–89 years old and four-fold greater risk than 70–to–79 years old. This result is consistent with Livingston and Bucher, 2020 and Onder et al. (2020), and assumes great importance, since the susceptibility to the COVID-19 is constant across all age groups (Istat-ISS, 2020). The prevalence and saturation indexes were highly and significantly correlated with the CFR in all the considered models

Table 4b
(Models 7–12). OLS regression at the provincial level between CFR and environmental, demographic, and healthcare factors.

Variables	Model 7	Model 8	Model 9	Model 10	Model 11	Model 12
Constant	31.4907***	41.079***	14.7429	18.1456	15.5861	18.8624
	[9.6085]	[11.7581]	[11.7796]	[11.3803]	[11.579]	[11.2617]
G. P.	−12.4582***	−13.0142***	−7.0346**	−7.6595**	−6.9201*	−7.578**
	[4.642]	[4.2388]	[3.4178]	[3.5788]	[3.5378]	[3.7071]
Temperature	−0.22	−0.4403*	−0.1162	−0.129	−0.2588	−0.2667
	[0.262]	[0.2471]	[0.2952]	[0.2905]	[0.2778]	[0.2745]
DTR	−2.7569***	−4.0316***	−1.7712*	−1.7872*	−1.8775*	−1.8874*
	[1.0324]	[1.2007]	[1.0379]	[1.0266]	[1.055]	[1.0434]
PM_{10} (μg/m^3)			0.0476	0.0454	0.0961	0.0939
			[0.0476]	[0.09]	[0.0989]	[0.0989]
Aged 70+	0.5006	0.6215*	0.64**		0.6583**	
	[0.3969]	[0.3324]	[0.3028]		[0.3031]	
Aged 80+				1.178**		1.2287**
				[0.5477]		[0.5514]
O_3 (> 120)	0.1084***					
	[0.0293]					
$PM_{2.5}$ (μg/m^3)		0.3127***				
		[0.1112]				
Altitude			−0.0002	−0.0005	−0.0002	−0.0005
			[0.0021]	[0.0021]	[0.0022]	[0.0022]
Prevalence			16.7325***	16.5201***		
			[3.7856]	[3.842]		
OB saturation					4.4258***	4.3711***
					[1.0094]	[1.0332]
Breusch-P p.	0.0003	0.0045	0.0000	0.0000	0.0003	0.0004
F-statistic	14.22***	10.48***	11.67***	11.77***	14.24***	14.14***
VIF	1.07–1.54	1.05–1.33	1.13–2.52	1.14–2.5	1.13–2.44	1.14–2.42
Observations	88	92	107	107	107	107
Adjusted R^2	0.4241	0.3531	0.4488	0.4492	0.4142	0.416

Notes: p, p-value. Standard errors (in brackets) are based on HC2 method developed by MacKinnon and White (1985). Significance level: p-value <0.01***; p-value <0.05**: p-value <0.1*.

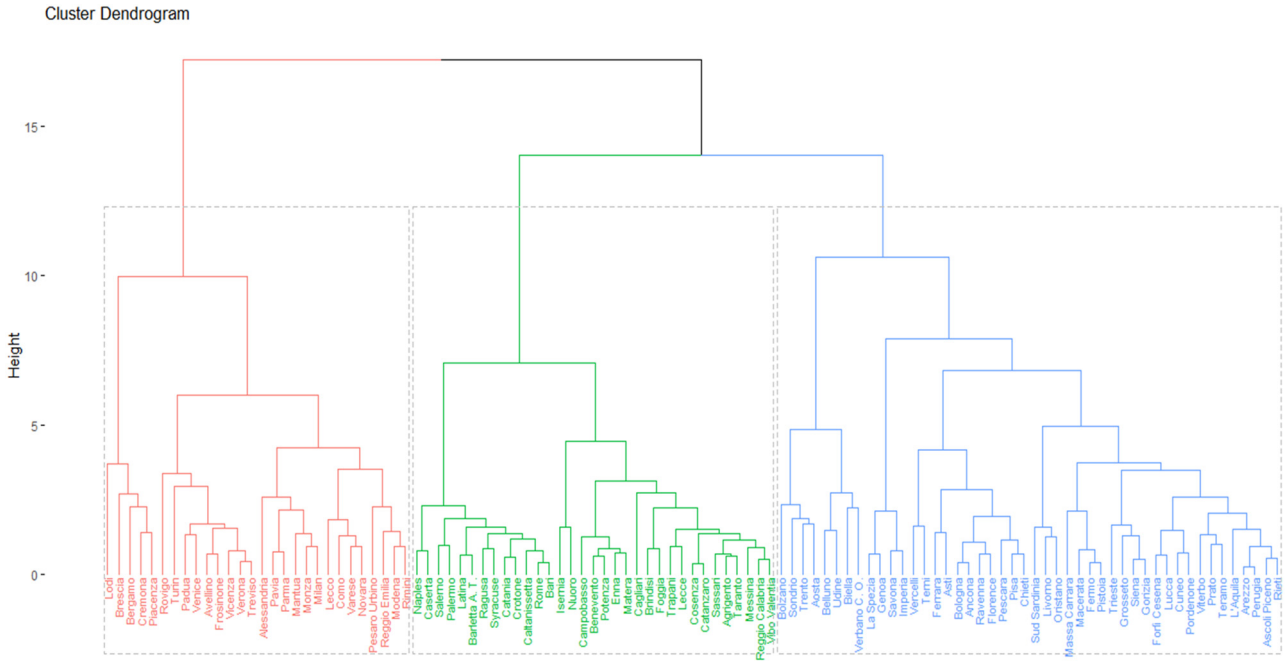


Fig. 3. Dendrogram of provinces obtained using Ward's method.

(9–12). Every 0.1 unit increase in COVID-19 prevalence and saturation of hospital ordinary beds was associated with an average change of 1.66% and 0.44% in the CFR, respectively. Moreover, the inclusion of prevalence and saturation in the models allowed an increase in the average r-square from 0.31 to 0.43. This is consistent with the previous findings.

The correlation matrix (Fig. S2) gave other interesting information, especially regarding air pollutants. Air pollutants concentrations were highly and positively correlated with each other, with a Pearson's r that ranged from 0.48 to 0.84. Moreover, consistent with Brandt et al. (2020), population density was moderately and positively correlated with air pollutants concentrations. Finally, as found in other studies (Bianconi et al., 2020; Pansini and Fornacca, 2020; Setti et al., 2020; Zhu et al., 2020), COVID-19 prevalence and air pollutant showed a good positive correlation, with a Pearson's r ranged from 0.36 to 0.65.

5.3. Cluster analysis

In Fig. 3, I presented the dendrogram obtained using Ward's method. The optimal number of clusters was identified by using two different methods: i) the EM (Expectation-Maximization) algorithm for Gaussian finite mixture model, proposed by Fraley et al. (2012) in the package 'Mclust' (R environment); and iii) the package 'NbClust' (R environment) that proposes 30 different indexes.[14] In the first case, according to BIC (Bayesian Information Criterion) score, the best model was VVI (varying volume, varying shape, and equal orientation) with 3 clusters. The majority of indices (7 out of 30) included in 'NbClust', also proposed 3 clusters. Therefore, by cutting the dendrogram at an approximately height of 13, I obtained three clusters with an increasing risk of mortality (Table 5). Each cluster was identified with a grey dotted rectangle (Fig. 3) and represented graphically through a map (Fig. 4). The cluster with the highest mortality risk (risk = 3) was composed by 25 provinces of northern Italy, 1 province of central Italy (Pesaro and Urbino), and 2 provinces of southern Italy (Frosinone and Avellino). The cluster with the medium mortality risk (risk = 2) was composed by 22 provinces of northern Italy, 18 provinces of central Italy, and 6 provinces of southern Italy (Chieti, L'Aquila, Oristano, Pescara, South Sardinia, and Teramo). Finally, the cluster with the lowest mortality risk (risk = 1) was composed by 2 provinces of central Italy (Latina and Rome), and 31 provinces of southern Italy.[15] Therefore, the highest mortality risk was concentrated in the north of the country, while the lowest risk was associated with southern provinces. Specifically, cluster 3 had an 8.62% higher CFR than cluster 1, and 5.3% higher CFR than cluster 2. Cluster 3 had a temperature of 2.58 degrees lower and a number of general practitioners per 1000 inhabitants of 0.24 lower than that for cluster 1. Moreover, cluster 3 had a PM_{10} of 11.23 $\mu g/m^3$ higher and a proportion of population aged 70 and above of 1.49% higher than that for cluster 1. Finally, the hospital bed saturation in cluster 3 was more than double than that for cluster 2, and 8 times greater than that for cluster 1. The results are consistent with Grasselli et al. (2020), according to which Lombardy's intensive care units already had an 85% to 90% occupancy ahead of the outbreak.[16]

6. Conclusions

To the best of my knowledge, this is one of the first studies to investigate the relationship between a wide set of heterogeneous factors and COVID-19 mortality. The OLS analysis showed that environmental, demographic, and healthcare factors played an important role in explaining the CFR variability. In particular, population aging, air pollutants concentrations (NO_2, O_3, PM_{10}, and $PM_{2.5}$), relative humidity,

[14] The outcomes of both methods are reported in the Supplementary Material (Figure S3 and Table S3).

[15] Some of these findings may be quite surprising. In fact, as pointed out by an anonymous referee the difference between Latina and Frosinone, for example, may be negligible and caused by environmental factors rather than COVID-19 prevalence or CFR. This is substantially true because in Frosinone the average concentrations of PM_{10} were 43% higher than those found in Latina. However, it should be considered that: i) Frosinone had a temperature of 2.3 degrees lower than Latina; and ii) above all, the prevalence/CFR data of COVID-19 used in this analysis were the only available official data for the considered period. For the rest, cluster analysis seemed to work well and confirmed the dichotomy between the north and the south of the country. In fact, not by change, despite the limited number of swabs at the peak of the epidemic, the average prevalence of COVID-19 in northern provinces was 0.2985, i.e. six times larger than that for southern provinces (0.0465).

[16] In fact, Lombardy was the hardest hit region on April 3–4, 2020, with 40% of the Italian confirmed cases.

Table 5
The clusters obtained by cutting dendrogram at an approximately height of 13.

Variables	Cluster (CL1) (Low risk)	Cluster (CL2) (Medium risk)	Cluster (CL3) (High risk)	CL3 - CL1
CFR	4.5758	7.8961	13.1957	8.6199
G. practitioners	1.0322	0.939	0.7947	−0.2375
Temperature	10.2758	7.1516	7.6946	−2.5812
Aged 70+	15.9355	19.3297	17.4285	1.493
PM_{10} ($\mu g/m^3$)	21.9211	21.9833	33.1464	11.2253
OB saturation	0.1393	0.5019	1.161	1.0217
Numerosity	33	46	28	–

COVID-19 prevalence, and critical care and ordinary beds saturation were positively correlated with the CFR. By the contrary, overall health care efficiency (IPS), physician density, and average temperature were negatively associated with CFR. Specifically, the inclusion of the COVID-19 prevalence and saturation indexes of ordinary and critical care beds explained up to 86% of the CFR variability. Therefore, a significant part of the CFR variability may have been caused by the massive stress of the Italian health system. The results are robust across several model specifications. Moreover, cluster analysis showed that the highest mortality risk was concentrated in northern Italy, while the lowest risk was associated with southern provinces.

However, this study also has some limitations that can be summarized as follows: i) first, a significant part of the patients died in hospital presented at least one comorbidity ahead of COVID-19 infection (ISS, 2020); ii) then, as pointed out in other studies (Spalt et al., 2016; Wang et al., 2019), the utilization of air pollution implies unavoidable measurement errors since most people usually stay indoors; iii) third, climatic variables, such as average temperature and DTR, refer just to the average historical values; and iv) finally, the data on prevalence are probably underestimated due to the limited number of swabs carried out in the early stages of the epidemic.

Finally, the study seemed to stress the importance of implementing quick and rational lockdown measures, of making patients comfortable, of implementing an action plan to discourage car use and decrease firm's pollution, and of buying ad hoc health care facilities, medical equipment, and devices to adequately tackle similar and unforeseeable emergencies.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

CRediT authorship contribution statement

Gaetano Perone: Writing - original draft, Writing - review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.scitotenv.2020.142523.

References

Baud, D., Qi, X., Nielsen-Saines, K., Musso, D., Pomar, L., Favre, G., 2020. Real estimates of mortality following COVID-19 infection. Lancet Infect. Dis. 20 (7), P773. https://doi.org/10.1016/S1473-3099(20)30195-X.

Bayer, C., Kuhn, M., 2020. Intergenerational ties and case fatality rates: a cross-country analysis. Institute of Labor Economic (IZA). Discussion Paper No. 13114, April 2020. https://covid-19.iza.org/wp-content/uploads/2020/04/dp13114.pdf.

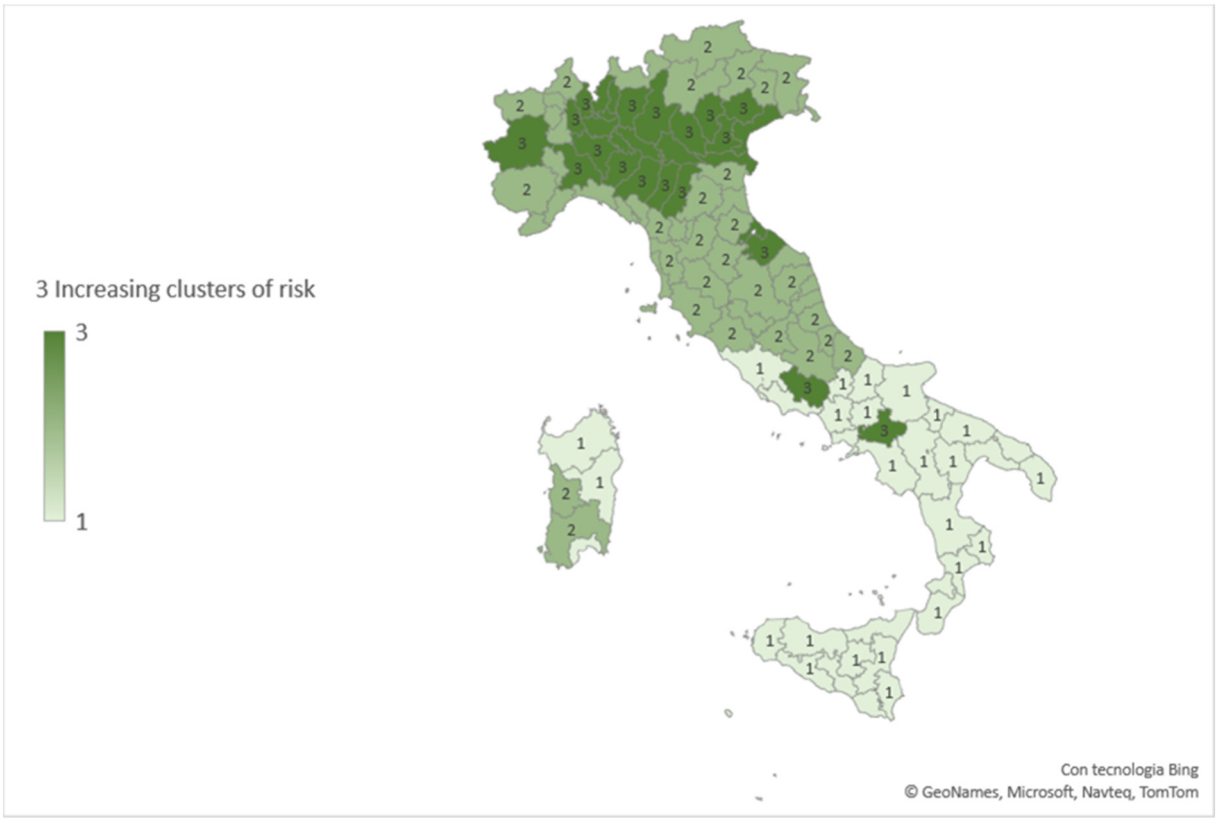


Fig. 4. Map of the 107 Italian provinces divided into three increasing clusters of risk.

Belloc, M., Buonanno, P., Drago, F., Galbiati, R., Pinotti, P., 2020. Cross-country Correlation Analysis for Research on COVID-19. Voxeu.org, March 28, 2020. https://voxeu.org/article/cross-country-correlation-analysis-research-covid-19.
Belsley, D.A., 1982. Assessing the presence of harmful collinearity and other forms of weak data through a test for signal-to-noise. J. Econ. 20 (2), 211–253. https://doi.org/10.1016/0304-4076(82)90020-3.
Belsley, D.E., Kuh, K., Welsch, R.E., 1980. Regression Diagnostics: Identifying Influential Data and Sources of Collinearity. Wiley and Sons, New York.
Bianconi, V., Bronzo, P., Banach, M., Sahebkar, A., Mannarino, M., Pirro, M., 2020. Particulate matter pollution and the COVID-19 outbreak: results from Italian regions and provinces. Arch. Med. Sci. 16 (1). https://doi.org/10.5114/aoms.2020.95336.
Bloomberg, 2019. These Are the World's Healthiest Nations. February 24. https://www.bloomberg.com/news/articles/2019-02-24/spain-tops-italy-as-world-s-healthiest-nation-while-u-s-slips.
Brandt, E.B., Beck, A.F., Mersha, T.B., 2020. Air pollution, racial disparities, and COVID-19 mortality. J. Allergy Clin. Immunol. 146 (1), 61–63. https://doi.org/10.1016/j.jaci.2020.04.035.
Breusch, T.S., Pagan, A.R., 1979. A simple test for heteroscedasticity and random coefficient variation. Econometrica: Journal of the Econometric Society 1287–1294.
Comunian, S., Dongo, D., Milani, C., Palestini, P., 2020. Air pollution and COVID-19: the role of particulate matter in the spread and increase of COVID-19's morbidity and mortality. International Journal of Environ. Res. and Public Health 17 (12), 4487. https://doi.org/10.3390/ijerph17124487.
Conticini, E., Frediani, B., Caro, D., 2020. Can atmospheric pollution be considered a co-factor in extremely high level of SARS-CoV-2 lethality in Northern Italy? Environ. Pollut. 261, 114465. https://doi.org/10.1016/j.envpol.2020.114465.
Demoskopika, 2018. La performance sanitaria. Indice di misurazione e valutazione dei sistemi regionali italiani https://www.corrieredellacalabria.it/wp-content/uploads/2018/03/IPS_Report_2018.pdf.
Demoskopika, 2019. La performance sanitaria. Indice di misurazione e valutazione dei sistemi regionali italiani https://www.quotidianosanita.it/allegati/allegato8783063.pdf.
Du, R.H., Liang, L.R., Yang, C.Q., Wang, W., Cao, T.Z., Li, M., ... Hu, M., 2020. Predictors of mortality for patients with COVID-19 pneumonia caused by SARS-CoV-2: a prospective cohort study. European Respiratory Journal 55 (5), 2000524. https://doi.org/10.1183/13993003.00524-2020.
European Commission, 2013. Urban-rural Typology, by NUTS3 Regions. https://ec.europa.eu/eurostat/cache/RCI/#?vis=urbanrural.urb_typology&lang=en.
Fraley, C., Raftery, A.E., Murphy, T.B., Scrucca, L., 2012. mclust version 4 for R: normal mixture modeling for model-based clustering, classification, and density estimation. University of Washington, Technical report No. 597 https://www.stat.washington.edu/sites/default/files/files/reports/2012/tr597.pdf.
Gan, W.Q., Davies, H.W., Koehoorn, M., Brauer, M., 2012. Association of long-term exposure to community noise and traffic-related air pollution with coronary heart disease mortality. Am. J. Epidemiol. 175 (9), 898–906. https://doi.org/10.1093/aje/kwr424.
GBD, 2017. Italy's health performance, 1990-2017: findings from the Global Burden of Disease Study 2017. Lancet Public Health 4 (12), E645–E657. https://doi.org/10.1016/S2468-2667(19)30189-6 Italy Collaborators (2019).
Ghasemi, A., Zahediasl, S., 2012. Normality tests for statistical analysis: a guide for non-statisticians. Int. J. Endocrinol. Metab. 10 (2), 486–489. https://doi.org/10.5812/ijem.3505.
Grasselli, G., Pesenti, A., Cecconi, M., 2020. Critical care utilization for the COVID-19 outbreak in Lombardy, Italy: early experience and forecast during an emergency response. JAMA 2020. https://doi.org/10.1001/jama.2020.4031.
Guan, W.J., Liang, W.H., Zhao, Y., Liang, H.R., Chen, Z.S., Li, Y.M., ... Ou, C.Q., 2020. Comorbidity and its impact on 1590 patients with Covid-19 in China: A nationwide Analysis. European Respiratory Journal 55 (5), 2000547. https://doi.org/10.1183/13993003.00547-2020.
Hair Jr., J.F., Anderson, R.E., Tatham, R.L., Black, W.C., 1995. Multivariate Data Analysis. 3th ed. New York, Macmillan.
Hamidi, S., Sabouri, S., Ewing, R., 2020. Does density aggravate the COVID-19 pandemic? J. Am. Plan. Assoc. https://doi.org/10.1080/01944363.2020.1777891.
I.Stat (database), d. http://dati.istat.it/.
Il Sole 24 Ore, 2019. Indice della Salute. https://lab24.ilsole24ore.com/indice-della-salute/indexT.php.
Ioannidis, J.P., Axfors, C., Contopoulos-Ioannidis, D.G., 2020. Population-level COVID-19 mortality risk for non-elderly individuals overall and for non-elderly individuals without underlying diseases in pandemic epicenters. Environ. Res. 188, 109890. https://doi.org/10.1016/j.envres.2020.109890.
ISS, Istituto Superiore di Sanità, 2020. Characteristics of COVID-19 Patients Dying in Italy Report Based on Available Data on March 26th, 2020. https://www.epicentro.iss.it/coronavirus/bollettino/Report-COVID-2019_26_marzo_eng.pdf.
Istat, 2019. Tavole di dati. Ambiente Urbano. December 18, 2019. https://www.istat.it/it/archivio/236912.
Istat-ISS, 2020. Impatto dell'epidemia COVID-19 sulla mortalità totale della popolazione residente primo trimestre 2020. May 4, 2020. https://www.istat.it/it/files/2020/05/Rapporto_Istat_ISS.pdf.
Italian Ministry of Health, d. www.salute.gov.it.
Jordan, R.E., Adab, P., Cheng, K.K., 2020. Covid-19: risk factors for severe disease and death. Br. Med. J. 368, m1198.
Livingston, E., Bucher, K., 2020. Coronavirus disease 2019 (COVID-19) in Italy. JAMA 323 (14), 1335. https://doi.org/10.1001/jama.2020.4344.
Long, J.S., Ervin, L.H., 2000. Using heteroscedasticity consistent standard errors in the linear regression model. The American Statistician 54 (3), 217–224. https://doi.org/10.1080/00031305.2000.10474549.
Ma, Y., Zhao, Y., Liu, J., He, X., Wang, B., Fu, S., ... Luo, B., 2020. Effects of temperature variation and humidity on the death of COVID-19 in Wuhan, China. Science of The Total Environment 724, 138226. https://doi.org/10.1016/j.scitotenv.2020.138226.
MacKinnon, J.G., White, H., 1985. Some heteroskedasticity-consistent covariance matrix estimators with improved finite sample properties. J. Econ. 29 (3), 305–325. https://doi.org/10.1016/0304-4076(85)90158-7.
Mipaaf, Ministero delle politiche agricole e forestali, 2019a. Osservatorio agroclimatico, statistiche regionali. https://www.politicheagricole.it/flex/FixedPages/Common/miepfy700_regioni.php/L/IT.
Mipaaf, Ministero delle politiche agricole e forestali, 2019b. Osservatorio agroclimatico, statistiche provinciali. https://www.politicheagricole.it/flex/FixedPages/Common/miepfy700_provincie.php/L/IT.
Mur, J., Lauridsen, J., 2007. Outliers and spatial dependence in cross-sectional regressions. Environ Plan A 39 (7), 1752–1769. https://doi.org/10.1068/a38207.
Ogen, Y., 2020. Assessing nitrogen dioxide (NO2) levels as a contributing factor to the coronavirus (COVID-19) fatality rate. Sci. Total Environ. 726, 138605. https://doi.org/10.1016/j.scitotenv.2020.138605.
Onder, G., Rezza, G., Brusaferro, S., 2020. Case-fatality rate and characteristics of patients dying in relation to COVID-19 in Italy. JAMA 323 (18), 1775–1776. https://doi.org/10.1001/jama.2020.4683.
Pansini, R., Fornacca, D., 2020. COVID-19 higher induced mortality in Chinese regions with lower air quality. medRxiv https://doi.org/10.1101/2020.05.28.20115832.
Rahman, M., Islam, M., Shimanto, M.H., Ferdous, J., Rahman, A.A.S., Sagor, P.S., Chowdhury, T., 2020. Temperature Extreme May Exaggerate the Mortality Risk of COVID-19 in the Low- and Middle-income Countries: A Global Analysis. https://doi.org/10.20944/preprints202006.0369.v1 (Preprints, 2020060369).
Sannigrahi, S., Pilla, F., Basu, B., Basu, A.S., 2020. The overall mortality caused by covid-19 in the european region is highly associated with demographic composition: a spatial regression-based approach. arXiv (preprint arXiv:2005.04029).
Scheiner, S., Ukaj, N., Hellmich, C., 2020. Mathematical modeling of COVID-19 fatality trends: death kinetics law versus infection-to-death delay rule. Chaos, Solitons Fractals 136, 109891. https://doi.org/10.1016/j.chaos.2020.109891.
Setti, L., Passarini, F., De Gennaro, G., Barbieri, P., Perrone, M.G., Borelli, M., ... Clemente, L., 2020. SARS-Cov-2RNA Found on Particulate Matter of Bergamo in Northern Italy: First Evidence. Environmental Research 188, 109754. https://doi.org/10.1016/j.envres.2020.109754.
Shapiro, S.S., Wilk, M.B., 1965. An analysis of variance test for normality (complete samples). Biometrika 52 (3/4), 591–611.
Spalt, E.W., Curl, C.L., Allen, R.W., Cohen, M., Adar, S.D., Stukovsky, K.H., ... Kaufman, J.D., 2016. Time–location patterns of a diverse population of older adults: the Multi-Ethnic Study of Atherosclerosis and Air Pollution (MESA Air). J. Expo. Sci. Environ. Epidemiol. 26 (4), 349–355. https://doi.org/10.1038/jes.2015.29.
Terna Group, 2019. Dati storici. https://download.terna.it/terna/8-DATI%20STORICI_8d736bdf116e207.pdf.
Tian, He-Zhong, Hao, Ji-Ming, Hu, Man-Yin, Nie, Yong-Feng, 2007. Recent trends of energy consumption and air pollution in China. J. Energy Eng. 133 (1), 4–12. https://doi.org/10.1061/(ASCE)0733-9402(2007)133:1(4).
Verity, R., Okell, L.C., Dorigatti, I., Winskill, P., Whittaker, C., Imai, N., ... Dighe, A., 2020. Estimates of the severity of coronavirus disease 2019: a model-based analysis. The Lancet infectious diseases 20 (6), 669–677. https://doi.org/10.1016/S1473-3099(20)30243-7.
Wang, M., Aaron, C.P., Madrigano, J., Hoffman, E.A., Angelini, E., Yang, J., ... Sheppard, L.E., 2019. Association between long-term exposure to ambient air pollution and change in quantitatively assessed emphysema and lung function. JAMA 322 (6), 546–556. https://doi.org/10.1001/jama.2019.10255.
Wang, D., Hu, B., Hu, C., Zhu, F., Liu, X., ... Zhiyong, P., 2020. Clinical characteristics of 138 hospitalized patients with 2019 Novel Coronavirus–Infected pneumonia in Wuhan, China. JAMA 323 (11), 1061–1069. https://doi.org/10.1001/jama.2020.1585.
Ward Jr., J.H., 1963. Hierarchical grouping to optimize an objective function. J. Am. Stat. Assoc. 58 (301), 236–244. https://doi.org/10.1080/01621459.1963.10500845.
Wooldridge, J.M., 2015. Introductory Econometrics. A Modern Approach. 5th edition. Mason, OH, South-Western.
World Health Organization, 2010. World Health Report 2010: Health Systems Financing – The Path to Universal Coverage. Geneva. https://www.who.int/whr/2010/whr10_en.pdf.
World Health Organization, 2020. https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19—11-march-2020.
Worldometers, 2020. https://www.worldometers.info/coronavirus/.
Wu, Z., McGoogan, J.M., 2020. Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China: summary of a report of 72314 cases from the Chinese Center for Disease Control and Prevention. JAMA 323 (13), 1239–1242. https://doi.org/10.1001/jama.2020.2648.
Wu, X., Nethery, R.C., Sabath, B.M., Braun, D., Dominici, F., 2020a. Exposure to air pollution and COVID-19 mortality in the United States. medRxiv https://doi.org/10.1101/2020.04.05.20054502.
Wu, Y., Jing, W., Liu, J., Ma, Q., Yuan, J., Wang, Y., ... Liu, M., 2020b. Effects of temperature and humidity on the daily new cases and new deaths of COVID-19 in 166 countries. Science of the Total Environment 729, 139051. https://doi.org/10.1016/j.scitotenv.2020.139051.
Yao, Y., Pan, J., Wang, W., Liu, Z., Kan, H., Qiu, Y., Meng, X., Wang, W., 2020. Association of particulate matter pollution and case fatality rate of COVID-19 in 49 Chinese cities. Sci. Total Environ. 741, 140396. https://doi.org/10.1016/j.scitotenv.2020.140396.
Zhu, J., Xie, J., Huang, F., Cao, L., 2020. Association between short-term exposure to air pollution and COVID-19 infection: evidence from China. Sci. Total Environ. 727, 138704. https://doi.org/10.1016/j.scitotenv.2020.138704.